

Representations of Genetic Tables, Bimagic Squares, Hamming Distances and Shannon Entropy

Inder J. Taneja

Departamento de Matemática
Universidade Federal de Santa Catarina
88.040-900 Florianópolis, SC, Brazil.
e-mail: ijtaneja@gmail.com
<http://www.mtm.ufsc.br/~taneja>

Abstract

In this paper we have established relations of the genetic tables with magic and bimagic squares. Connections with Hamming distances, binomial coefficients are established. The idea of Gray code is applied. Shannon entropy of magic squares of order 4×4 , 8×8 and 16×16 are also calculated. Some comparison is also made. Symmetry among restriction enzymes having four letters is also studied.

Key words: Genetic Code, Codon, Magic Squares, Hamming distances, Probability distributions, Shannon entropy.

1 Introduction

Genetic code is the set of rules by which information encoded in RNA/DNA is translated into amino acid sequences in living cells. The bases for the encoded information are nucleotides. There are four nucleotide bases for RNA: Adenine, Uracil, Guanine, and Cytosine, which are labeled by A , U , G and C respectively, (in DNA Uracil is replaced by Thymine (T)). In canonical genetic code, codons are tri-nucleotide sequences such that each triplet relates to an amino acid. For example, the codon CAG encodes the amino acid Glutamine. Amino acids are the basic building blocks of proteins. It stimulated interest of other researchers to study how genetic code was translated into amino acids. There are 20 different amino acids (plus start and stop codons), and since there are four nucleotide bases, A , U , T and C , there are 4^n different combinations of bases, for a string of length n . Therefore, $n = 3$ is the smallest number of bases that could be used to represent the 20 different amino acids. There is degeneracy between the codons, i.e., more than one codon can represent the same amino acid; however, two different amino acids cannot be represented by the same codon. The following CODON table is well-known in the literature.

	T	C	A	G	
T	TTT (<i>Phe</i>)	TCT (<i>Ser</i>)	TAT (<i>Tyr</i>)	TGT (<i>Cys</i>)	T
	TTC (<i>Phe</i>)	TCC (<i>Ser</i>)	TAC (<i>Tyr</i>)	TGC (<i>Cys</i>)	C
	TTA (<i>Leu</i>)	TCA (<i>Ser</i>)	TAA (<i>Stop</i>)	TGA (<i>Stop</i>)	A
	TTG (<i>Leu</i>)	TCG (<i>Ser</i>)	TAG (<i>Stop</i>)	TGG (<i>Trp</i>)	G
C	CTT (<i>Leu</i>)	CCT (<i>Pro</i>)	CAT (<i>His</i>)	CGT (<i>Arg</i>)	T
	CTC (<i>Leu</i>)	CCC (<i>Pro</i>)	CAC (<i>His</i>)	CGC (<i>Arg</i>)	C
	CTA (<i>Leu</i>)	CCA (<i>Pro</i>)	CAA (<i>Glu</i>)	CGA (<i>Arg</i>)	A
	CTG (<i>Leu</i>)	CCG (<i>Pro</i>)	CAG (<i>Glu</i>)	CGG (<i>Arg</i>)	G
A	ATT (<i>Ile</i>)	ACT (<i>Thr</i>)	AAT (<i>Asn</i>)	AGT (<i>Ser</i>)	T
	ATC (<i>Ile</i>)	ACC (<i>Thr</i>)	AAC (<i>Asn</i>)	AGC (<i>Ser</i>)	C
	ATA (<i>Ile</i>)	ACA (<i>Thr</i>)	AAA (<i>Lys</i>)	AGA (<i>Arg</i>)	A
	ATG (<i>Met</i>)	ACG (<i>Thr</i>)	AAG (<i>Lys</i>)	AGG (<i>Arg</i>)	G
G	GTT (<i>Val</i>)	GCT (<i>Ala</i>)	GAT (<i>Asp</i>)	GGT (<i>Gly</i>)	T
	GTC (<i>Val</i>)	GCC (<i>Ala</i>)	GAC (<i>Asp</i>)	GGC (<i>Gly</i>)	C
	GTA (<i>Val</i>)	GCA (<i>Ala</i>)	GAA (<i>Glu</i>)	GGA (<i>Gly</i>)	A
	GTG (<i>Val</i>)	GCG (<i>Ala</i>)	GAG (<i>Glu</i>)	GGG (<i>Gly</i>)	G

The DNA (Deoxyribonucleic acid) molecule residing in the cell nucleus encodes information conventionally represented as a symbolic string over the alphabet. The combination between single strands of DNA takes

place according to “Watson-Crick [12] complementarity” that says that the only permissible combinations between bases are $A - T$ or $T - A$ and $C - G$ or $G - C$ hence one strand can easily be used to predict the other in a double stranded chain. Let us consider the following configurations of 4^n for each value of n .

- (i) For $n = 1$: In this case we have $4^1 = 4$. This gives

$$M_1 := \begin{bmatrix} C & A \\ T & G \end{bmatrix}.$$

- (ii) For $n = 2$: In this case we have $4^2 = 16$. This gives

$$M_2 := \begin{bmatrix} CC & AC & TC & GC \\ CA & AA & TA & GA \\ CT & AT & TT & GT \\ CG & AG & TG & GG \end{bmatrix}.$$

- (iii) For $n = 3$: In this case we have $4^3 = 64$. This gives

$$M_3 := \begin{bmatrix} CCC & ACC & TCC & GCC & CTC & ATC & TTC & GTC \\ CCA & ACA & TCA & GCA & CTA & ATA & TTA & GTA \\ CCT & ACT & TCT & GCT & CTT & ATT & TTT & GTT \\ CCG & ACG & TCG & GCG & CTG & ATG & TTG & GTG \\ CAC & AAC & TAC & GAC & CGC & AGC & TGC & GGC \\ CAA & AAA & TAA & GAA & CGA & AGA & TGA & GGA \\ CAT & AAT & TAT & GAT & CGT & AGT & TGT & GGT \\ CAG & AAG & TAG & GAG & CGG & AGG & TGG & GGG \end{bmatrix}.$$

- (iv) For $n = 4$, we have M_4 with $4^4 = 256$ combinations of blocks of four letters, for $n = 5$, we have M_5 with $4^5 = 1024$, etc.

2 Gray Codes: Binary representations

2.1 First Approach

Let us represent the letters C, A, T and G in two different ways:

- (i) $C = 00$, $A = 01$, $T = 10$ and $G = 11$.

- (ii) $C = 1$, $A = 2$, $T = 3$ and $G = 4$.

- (iii) The CODON table given above is formed of three letters out of four, i.e., A, T, G and C . According to (i), we can write, for example, $TTA \sim 101001$, $AGC \sim 011100$, etc. Thus we have six digit binary representations of 64 members available in CODON table. Let us apply the change of base 2 to base 10 (decimal) using the formula $(abcdef)_2 := a \cdot 2^5 + b \cdot 2^4 + c \cdot 2^3 + d \cdot 2^2 + e \cdot 2^1 + f \cdot 2^0$ and then writing $(abcdef)_2 + 1$, we have, $TTA \sim 101001 \sim 41$ and $AGC \sim 011100 \sim 28$, etc. Similarly, we can write the four digits binary representation in decimal forms, such as $(abcd)_2 := a \cdot 2^3 + b \cdot 2^2 + c \cdot 2^1 + d \cdot 2^0$, and then writing $(abcd)_2 + 1$. Just for simplicity, we have added 1.

The notations given in (i) and (ii) can be seen in [4, 6, 7, 8, 10]. Decimal representation of the numbers is given by $C = (00)_2 \sim 0$, $A = (01)_2 \sim 1$, $T = (10)_2 \sim 2$ and $G = (11)_2 \sim 3$. For simplicity, we have added 1 and considered in (ii) as 1, 2, 3, and 4 instead of 0, 1, 2 and 3 respectively. We shall use frequently these three

representations and shall bring magic squares of different orders According to above notations we have

$$M_1 := \begin{bmatrix} C & A \\ T & G \end{bmatrix} \sim \begin{bmatrix} 00 & 01 \\ 10 & 11 \end{bmatrix} \sim \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \quad (1)$$

$$M_2 := \begin{bmatrix} 0000 & 0100 & 1000 & 1100 \\ 0001 & 0101 & 1001 & 1101 \\ 0010 & 0110 & 1010 & 1110 \\ 0011 & 0111 & 1011 & 1111 \end{bmatrix}, \quad (2)$$

$$M_2 := \begin{bmatrix} 11 & 21 & 31 & 41 \\ 12 & 22 & 32 & 42 \\ 13 & 23 & 33 & 43 \\ 14 & 24 & 34 & 44 \end{bmatrix} \quad (3)$$

and

$$M_2 := \begin{bmatrix} 1 & 5 & 9 & 13 \\ 2 & 6 & 10 & 14 \\ 3 & 7 & 11 & 15 \\ 4 & 8 & 12 & 16 \end{bmatrix}. \quad (4)$$

The expressions appearing in (1) are due to (i), (ii) and (iii). The expression (2) is due to (i), (3) is due to (ii) and (3) is due to (iii). Similar tables can also be written for the matrix M_3 Some of them can be seen in [4, 6, 7, 8, 10].

2.2 Second Approach

Following [2, 3], we use the following correspondence for the nucleotides and two-bit Gray codes:

$$C \sim \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad A \sim \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad T \sim \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ and } G \sim \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

The genetic code-based matrix, which will contain all nucleotide strings of length n is defined as M_n . The Gray code sequences represented by M_n will be denoted by a $2^n \times 2^n$ matrix. Here are corresponding Gray code representations

$$M_1 := \begin{bmatrix} C & A \\ T & G \end{bmatrix} \sim \begin{bmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{bmatrix} \sim \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

In information theory, the *Hamming distance* between two strings of equal length is the number of positions for which the corresponding symbols are different. Put another way, it measures the minimum number of *substitutions* required to change one into the other, or the number of errors that transformed one string into the other. Thus we observe that the *Hamming distances* of letters C and G is 0 and of letters A and T is 1. Replacing the same in the other cases we have

$$M_2 := \begin{bmatrix} CC & AC & TC & GC \\ CA & AA & TA & GA \\ CT & AT & TT & GT \\ CG & AG & TG & GG \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix}$$

and

$$M_3 := \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 & 2 & 3 & 3 & 2 \\ 1 & 2 & 2 & 1 & 2 & 3 & 3 & 2 \\ 1 & 1 & 1 & 0 & 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 & 0 & 1 & 1 & 0 \\ 2 & 3 & 3 & 2 & 1 & 2 & 2 & 1 \\ 2 & 3 & 3 & 2 & 1 & 2 & 2 & 1 \\ 1 & 2 & 2 & 1 & 0 & 1 & 1 & 0 \end{bmatrix}.$$

In the theory of discrete signal processing as a fundamental operation for binary variables, modulo-2 addition is utilized broadly. By definition, the modulo-2 addition of two numbers written in binary notation is made in a bitwise manner in accordance with the following rules:

$$0 + 0 = 0, \quad 1 + 0 = 1, \quad 0 + 1 = 1, \quad 1 + 1 = 0$$

For example, modulo-2 addition of two binary numbers 110 and 101, gives the result $110 \oplus 101 = 011(3)$, where 3 is the decimal representation of 011. In case of 10 and 01, we have $10 \oplus 01 = 11(3)$, where 3 is the decimal representation of 11 (\oplus is the symbol for modulo-2 addition). The distance in this symmetry group is known as the *Hamming distance*. The modulo-2 addition of any two binary numbers always results in a new number from the same series. If any system of elements demonstrates its connection with diadic shifts, it indicates that the structural organization of its system is related to the logic of modulo-2 addition. In particular we have

$$M_1 := \begin{bmatrix} C & A \\ T & G \end{bmatrix} \sim \left[\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right] \sim \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$M_2 := \begin{bmatrix} CC & AC & TC & GC \\ CA & AA & TA & GA \\ CT & AT & TT & GT \\ CG & AG & TG & GG \end{bmatrix} \sim \begin{pmatrix} 00 & 10 & 10 & 00 \\ 01 & 11 & 11 & 01 \\ 01 & 11 & 11 & 01 \\ 00 & 10 & 10 & 00 \end{pmatrix}$$

and

$$M_3 := \begin{bmatrix} 000 & 100 & 100 & 000 & 010 & 110 & 110 & 010 \\ 001 & 101 & 101 & 001 & 011 & 111 & 111 & 011 \\ 001 & 101 & 101 & 001 & 011 & 111 & 111 & 011 \\ 000 & 100 & 100 & 000 & 010 & 110 & 110 & 010 \\ 010 & 110 & 110 & 010 & 000 & 100 & 100 & 000 \\ 011 & 111 & 111 & 011 & 001 & 101 & 101 & 001 \\ 011 & 111 & 111 & 011 & 001 & 101 & 101 & 001 \\ 010 & 110 & 110 & 010 & 000 & 100 & 100 & 000 \end{bmatrix}.$$

The results are obtained by using the binary operations given above for example, $GT \sim \begin{pmatrix} 11 \\ 10 \end{pmatrix} \sim 01$, i.e.,

$11 \oplus 10 = 01$, $ACT \sim \begin{pmatrix} 001 \\ 100 \end{pmatrix} \sim 101$, i.e., $001 \oplus 100 = 101$, etc. The decimal transformations are

$$(00)_2 \sim 0, \quad (01)_2 \sim 1, \quad (10)_2 \sim 2, \quad (11)_2 \sim 3$$

and

$$\begin{aligned} (000)_2 &\sim 0, & (001)_2 &\sim 1, & (010)_2 &\sim 2, & (011)_2 &\sim 3 \\ (100)_2 &\sim 4 & (101)_2 &\sim 5, & (110)_2 &\sim 6, & (111)_2 &\sim 7 \end{aligned}$$

This gives

$$M_2 := \begin{bmatrix} 0 & 2 & 2 & 0 \\ 1 & 3 & 3 & 1 \\ 1 & 3 & 3 & 1 \\ 0 & 2 & 2 & 0 \end{bmatrix}$$

and

$$M_3 := \begin{bmatrix} 0 & 4 & 4 & 0 & 2 & 6 & 6 & 2 \\ 1 & 5 & 5 & 1 & 3 & 7 & 7 & 3 \\ 1 & 5 & 5 & 1 & 3 & 7 & 7 & 3 \\ 0 & 4 & 4 & 0 & 2 & 6 & 6 & 2 \\ 2 & 6 & 6 & 2 & 0 & 4 & 4 & 0 \\ 3 & 7 & 7 & 3 & 1 & 5 & 5 & 1 \\ 3 & 7 & 7 & 3 & 1 & 5 & 5 & 1 \\ 2 & 6 & 6 & 2 & 0 & 4 & 4 & 0 \end{bmatrix}.$$

3 Reconfiguration Tables and Magic Squares

This section deals with the reconfigurations of matrices given above. These reconfigurations are made in such a way that using the notations given in section 2.1, lead use to magic squares or bimagic squares. Here below are the definitions of magic and bimagic squares

- A **magic square** is a collection of numbers put as a square matrix, where the sum of element of each row, each column and two principal diagonals are the same sum. For simplicity, let us write it as **S1**
- **Bimagic square** is a magic square where the sum of squares of each element of rows, columns and two principal diagonals are the same. For simplicity, let us write it as **S2**.

3.1 Reconfiguration Tables of order 4x4

Let us reconsider the matrix M_2 as

$$M_2^{4 \times 4} := \begin{bmatrix} AT & TG & CC & GA \\ CA & GC & AG & TT \\ GG & CT & TA & AC \\ TC & AA & GT & CG \end{bmatrix}. \quad (5)$$

In the above configuration, we have permutations of letters C , A , T and G are the *first* and *second* places, in various situations, for example, in each row, in each column, main diagonals, each group of order 2×2 , middle group of order 2×2 , four corner elements, symmetrical diagonals, etc. These configurations are of the following type

$$\begin{bmatrix} AT & TG \\ CA & GC \end{bmatrix}, \begin{bmatrix} GC & AG \\ CT & TA \end{bmatrix}, [CA \quad GC \quad AG \quad TT], \begin{bmatrix} CC \\ AG \\ TA \\ GT \end{bmatrix}, \text{etc.}$$

Here below are 20 combinations where the first and second members are the permutations of the letters C , A , T and G :

1	2	3	4	5	5	5	5	9	9	10	10	14	15	15	14	17	19	20	18
1	2	3	4	6	6	6	6	9	9	10	10	16	13	13	16	19	17	18	20
1	2	3	4	7	7	7	7	11	11	12	12	16	13	13	16	20	18	17	19
1	2	3	4	8	8	8	8	11	11	12	12	14	15	15	14	18	20	19	17

According to notations (i), (ii) and (iii) given in section 2.1, we have

0110	1011	0000	1101
23	34	11	42
7	12	1	14
0001	1100	0111	1010
12	41	24	33
2	13	8	11
1111	0010	1001	0100
44	13	32	21
16	3	10	5
1000	0101	1110	0011
31	22	43	14
9	6	15	4

In all the three situations we have magic squares of order 4×4 with $S1^{4 \times 4} := 2222$, $S1^{4 \times 4} := 110$ and $S1^{4 \times 4} := 34$ respectively. The last one is well-known *Khajurao magic square of order* 4×4 :

7	12	1	14
2	13	8	11
16	3	10	5
9	6	15	4

The above magic square of order 4×4 is one of the *most perfect magic* square known in the literature. Its connections with genetic code can be seen in [11]. This is one of the very little work available on magic squares connecting DNA.

According to binary operations given in section 2.2, we have

$$M_B^{4 \times 4} := \begin{bmatrix} 11 & 10 & 00 & 01 \\ 01 & 00 & 10 & 11 \\ 00 & 01 & 11 & 10 \\ 10 & 11 & 01 & 00 \end{bmatrix} \sim \begin{bmatrix} 3 & 2 & 0 & 1 \\ 1 & 0 & 2 & 3 \\ 0 & 1 & 3 & 2 \\ 2 & 3 & 1 & 0 \end{bmatrix}.$$

We observe that the matrix $M_2^{4 \times 4}$ is a composition of two *mutually orthogonal diagonalize Latin squares*, while the matrix $M_B^{4 \times 4}$ is not a diagonalize Latin square.

3.2 Reconfiguration Tables of order 8x8

Here we shall reorganize the CODON table or the matrix M_3 given in section 1 in such way that it becomes as magic square of order 8×8 . We shall present two different ways:

1. Four *magic squares* of order 4×4 of the sum $S1^{4 \times 4}$ having all the properties of the configuration matrix $M_2^{4 \times 4}$ given by (5).
2. *Bimagic square* of order 8×8 .

3.2.1 First Representations

Let us consider the following reorganization of matrix M_3 or the above CODON table:

CCC	TAT	GTG	AGA	CAA	TCG	GGT	ATC
GTA	AGG	CCT	TAC	GGC	ATT	CAG	TCA
AGT	GTC	TAA	CCG	ATG	GGA	TCC	CAT
TAG	CCA	AGC	GTT	TCT	CAC	ATA	GGG
CTG	TGA	GCC	AAT	CGT	TTC	GAA	ACG
GCT	AAC	CTA	TGG	GAG	ACA	CGC	TTT
AAA	GCG	TGT	CTC	ACC	GAT	TTG	CGA
TGC	CTT	AAG	GCA	TTA	CGG	ACT	GAC

In the above configuration we have the same properties of the magic square of order 4×4 given by (5), i.e., there are many permutation of the letters *C*, *A*, *T* and *G* in the *first*, *second* and *third* places. Each block of order 4×4 , half-row, half-column, half-principle diagonals etc. are also follow the same property. There are much more combinations of this type in the above configuration. In another way we can say there is a uniform distributions of letters *C*, *A*, *T* and *G*. Using the notations (i), (ii) and (iii) given in section 2.1, we have

000000 111 1	100110 323 39	111011 434 60	011101 242 30	000101 122 6	100011 314 36	111110 443 63	011000 231 25
111001 432 58	011111 244 32	000010 113 3	100100 321 37	111100 441 61	011010 233 27	000111 124 8	100001 312 34
011110 243 31	111000 431 57	100101 322 38	000011 114 4	011011 234 28	111101 442 62	100000 311 33	000110 123 7
100111 324 40	000001 112 2	011100 241 29	111010 433 59	100010 313 35	000100 121 5	011001 232 26	111111 444 64
001011 134 12	101101 342 46	110000 411 49	010110 223 23	001110 143 15	101000 331 41	110101 422 54	010011 214 20
110010 413 51	010100 221 21	001001 132 10	101111 344 48	110111 424 56	010001 212 18	001100 141 13	101010 333 43
010101 222 22	110011 414 52	101110 343 47	001000 131 9	010000 211 17	110110 423 55	101011 334 44	001101 142 14
101100 341 45	001010 133 11	010111 224 24	110001 412 50	101001 332 42	001111 144 16	010010 213 19	110100 421 53

The above table brings three different *magic squares* of order 8×8 , i.e., in each case we have $S1^{8 \times 8} := 444444$, $S1^{8 \times 8} := 2220$ and $S1^{8 \times 8} := 260$ respectively. Moreover, the above magic square is also bimagic in columns, i.e., for each column we have $S2^{8 \times 8} := 44893328844$, $S2^{8 \times 8} := 717060$ and $S2^{8 \times 8} := 11180$ respectively. Also, each block of order 4×4 is a magic square with $S1^{4 \times 4} := 222222$. Sum of each block of order 2×2 also has the same sum as of $S1^{4 \times 4}$.

3.2.2 Second Representation

We observe that the above magic square is bimagic only in columns. Here below we shall present a little different representation of CODON table resulting in *bimagic square* of order 8×8 . Let us consider the following configuration:

CGG	TTC	TCG	CAC	ATT	GGA	GAT	ACA
ATA	GGT	GAA	ACT	CGC	TTG	TCC	CAG
CCC	TAG	TGC	CTG	AAA	GCT	GTA	AGT
AAT	GCA	GTT	AGA	CCG	TAC	TGG	CTC
TAA	CCT	CTA	TGT	GCC	AAG	AGC	GTG
GCG	AAC	AGG	GTC	TAT	CCA	CTT	TGA
TTT	CGA	CAT	TCA	GGG	ATC	ACG	GAC
GGC	ATG	ACC	GAG	TTA	CGT	CAA	TCT

In the above configuration we have permutations of four letters C , A , T and G only in the *first* and *third* place in each block of order 2×2 , half-row, half-column, half-principal diagonals etc. Again, using the notations (i), (ii) and (iii) given in section 2.1, we have

001111 144 16	101000 331 41	100011 314 36	000100 121 5	011010 233 27	111101 442 62	110110 423 55	010001 212 18
011001 232 26	111110 443 63	110101 422 54	010010 213 19	001100 141 13	101011 334 44	100000 311 33	000111 124 8
000000 111 1	100111 324 40	101100 341 45	001011 134 12	010101 222 22	110010 413 51	111001 432 58	011110 243 31
010110 223 23	110001 412 50	111010 433 59	011101 242 30	000011 114 4	100100 321 37	101111 344 48	001000 131 9
100101 322 38	000010 113 3	001001 132 10	101110 343 47	110000 411 49	010111 224 24	011100 241 29	111011 434 60
110011 414 52	010100 221 21	011111 244 32	111000 431 57	100110 323 39	000001 112 2	001010 133 11	101101 342 46
101010 333 43	001101 142 14	000110 123 7	100001 312 34	111111 444 64	011000 231 25	010011 214 20	110100 421 53
111100 441 61	011011 234 28	010000 211 17	110111 424 56	101001 332 42	001110 143 15	000101 122 6	100010 313 35

The above table bring us three *bimagic squares* of order 8×8 with the following sums:

- (a) $S1^{8 \times 8} := 444444 = 12012 \times 37$; $S2^{8 \times 8} := 44893328844 = 1213333212 \times 37$
- (b) $S1^{8 \times 8} := 2220 = 60 \times 37$; $S2^{8 \times 8} := 717060 = 19380 \times 37$
- (c) $S1^{8 \times 8} := 260$; $S2^{8 \times 8} := 11180$.

We observe that (a) and (b) both $S1^{8 \times 8}$ and $S2^{8 \times 8}$ are multiple of 37. This is not true in case of (c). Still, in (a), (b) and (c) each block of order 2×4 is also *bimagic*

3.3 Connections with Prime Number 37

Many authors [5, 6, 9] made connections with *prime number 37*. Here also we shall bring some interesting connections with this prime number. Both the representations we have three cases. In the first, we can write the sum $S1^{8 \times 8} := 444444 = 12 \times 1001 \times 37$. In the second case we have $S1^{8 \times 8} := 2220 = 12 \times 5 \times 37$. Also, in both these cases we have half-sum of rows, columns and two principal diagonal as multiple of 37, i.e., in the first case we have $\frac{1}{2}S^{8 \times 8} = 222222 = 6 \times 1001 \times 37$, and in the second case we have $\frac{1}{2}S^{8 \times 8} = 1110 = 30 \times 37$. There are many other combinations in the above table giving connection with 37. For example each block of 2×2 is of sum $\frac{1}{2}S1^{8 \times 8}$. In case of $S2^{8 \times 8}$ we have $S2^{8 \times 8} := 44893328844 = 1213333212 \times 37$ and $S2^{8 \times 8} := 717060 = 37 \times 19380$. In the first representation, we have *bimagic* sum only in each column.

3.4 Hamming Distances and Binomial Coefficients

The idea of *Hamming distances* is given in section 2.2. Here we consider more representations to bring *binomial coefficients* and *bimagic squares*. Kappraff and Adamson [4] considered $C = G$ and $A = U/T$. Following the idea of Watson and Crick [12], they [4] considered it as $A = T = 2$ and $C = G = 3$. For simplicity,

let us consider here $A = T = a$ and $C = G = b$, where it is understood that $TTG = a \times a \times b = a^2b$, $AGC = a \times b \times b = ab^2$, etc. Accordingly, we have

(i) For $n = 1$:

0	1
b	a
1	0
a	b

(ii) For $n = 2$:

1	0	2	1
ab	b^2	a^2	ab
2	1	1	0
a^2	ab	ab	b^2
0	1	1	2
b^2	ab	ab	a^2
1	2	0	1
ab	a^2	b^2	ab

(iii) For $n = 3$: According to configuration given in section 3.2.1, we have

0	3	1	2	2	1	1	2
a^3	b^3	ab^2	a^2b	a^2b	ab^2	ab^2	a^2b
2	1	1	2	0	3	1	2
a^2b	ab^2	ab^2	a^2b	a^3	b^3	ab^2	a^2b
2	1	3	0	2	1	1	2
a^2b	ab^2	b^3	a^3	a^2b	ab^2	ab^2	a^2b
2	1	1	2	2	1	3	0
a^2b	ab^2	ab^2	a^2b	a^2b	ab^2	b^3	a^3
1	2	0	3	1	2	2	1
ab^2	a^2b	a^3	b^3	ab^2	a^2b	a^2b	ab^2
1	2	2	1	1	2	0	3
ab^2	a^2b	a^2b	ab^2	ab^2	a^2b	a^3	b^3
3	0	2	1	1	2	2	1
b^3	a^3	a^2b	ab^2	ab^2	a^2b	a^2b	ab^2
1	2	2	1	3	0	2	1
ab^2	a^2b	a^2b	ab^2	b^3	a^3	a^2b	ab^2

According to configuration given in section 3.2.2 we have

1 ab^2	2 a^2b	0 b^3	1 ab^2	2 a^2b	3 a^3	1 ab^2	2 ab^2
0 b^3	1 ab^2	1 ab^2	2 a^2b	1 a^2b	2 a^2b	2 a^2b	3 a^3
0 b^3	1 ab^2	1 ab^2	2 a^2b	1 ab^2	2 a^2b	2 a^2b	3 a^3
1 ab^2	2 a^2b	0 b^3	1 ab^2	2 a^2b	3 a^3	1 ab^2	2 a^2b
2 a^2b	1 ab^2	1 ab^2	0 b^3	3 a^3	2 a^2b	2 a^2b	1 ab^2
1 ab^2	0 b^3	2 a^2b	1 ab^2	2 a^2b	1 ab^2	3 a^3	2 a^2b
1 ab^2	0 b^3	2 a^2b	1 ab^2	2 a^2b	1 ab^2	3 a^3	2 a^2b
2 a^2b	1 ab^2	1 ab^2	0 b^3	3 a^3	2 a^2b	2 a^2b	1 ab^2

The interesting fact in the above tables is that in the first case, it is symmetric in rows, columns and principal diagonals, while it is not true in the second case. In the second case it holds only in rows. The tables studied above gives us the following frequency distributions:

n	Hamming distances	Frequency distributions	Binomial coefficients	Sum
1	0 1	$2^1 = 2$	$a \ b$	$(a + b)^1$
2	0 1 2	$2^2 = 4$	$a^2 \ 2ab \ b^2$	$(a + b)^2$
3	0 1 2 3	$2^3 = 8$	$a^3 \ 3a^2b \ 3ab^2 \ b^3$	$(a + b)^3$

For more properties of above table refer to [4].

3.5 Binary Operations

Considering the notations and binary operations given in section 2.2, i.e., $CGT := \begin{pmatrix} 011 \\ 010 \end{pmatrix} \sim 001$, $ATC := \begin{pmatrix} 010 \\ 100 \end{pmatrix} \sim 110$, etc. This operations gives us eight possibilities, i.e., 000, 001, 010, 011, 100, 101, 110 and 111. Let us represent $000 \rightarrow a$, $001 \rightarrow b$, $010 \rightarrow c$, $011 \rightarrow d$, $100 \rightarrow e$, $101 \rightarrow f$, $110 \rightarrow g$ and $111 \rightarrow h$. Instead of decimal representations as 0, 1, 2, 3, 4, 5, 6 and 7 we have considered here the letters a, b, c, d, e, f, g and h respectively. According to section 3.2.1, we have the following table:

a	h	c	f	d	e	b	g
d	e	b	g	a	h	c	f
f	c	h	a	g	b	e	d
g	b	e	d	f	c	h	a
c	f	a	h	b	g	d	e
b	g	d	e	c	f	a	h
h	a	f	c	e	d	g	b
e	d	g	b	h	a	f	c

We observe that we have 16 matrices of order 2×2 divided in two groups formed by the elements $(a, d, e, h) \sim (000, 011, 100, 111)$ and $(b, c, f, g) \sim (001, 010, 101, 110)$. The above configuration is well-known **diagonalize Latin square of order 8×8** In the second case, i.e., for the section 3.2.2, we don't have symmetric configuration. See below:

a	g	e	c	h	b	d	f
h	b	d	f	a	g	e	c
a	g	e	c	h	b	d	f
h	b	d	f	a	g	e	c
h	b	d	f	a	g	e	c
a	g	e	c	h	b	d	f
h	b	d	f	a	g	e	c
a	g	e	c	h	b	d	f

4 Restriction Enzymes

There are (ref. Reiner [11]) 402 known restriction enzymes. Out of these 402 enzymes, 108 cut at a tetrameric sequence containing the four different bases. Of these 108, 100% have either *AT* or *GC* dimers (or both) in the sequence. None of 108 enzymes have *G* apart from *C* and *A* apart from *T*, as in *AGTC*. Thus, all 108 enzymes cut at the tetrameric sequence which is complementary to its reverse cyclic permutation. The specific antiparallel sequences and the enzymes at which they cut are listed below. Reiner [11] considered following two different combinations of four letters having together, either *AT* – *TA* or *GC* – *CG* specifying antiparallel enzymes See the table below:

Antiparallel <i>A</i> – <i>T</i> , <i>G</i> – <i>C</i> in same orientation (88)	Antiparallel <i>A</i> – <i>T</i> , <i>G</i> – <i>C</i> in opposite orientation (20)
<i>AGCT</i> (9)	<i>TAGC</i> (0)
<i>CGTA</i> (0)	<i>ACGT</i> (2)
<i>TACG</i> (0)	<i>GTAC</i> (4)
<i>CTAG</i> (9)	<i>GCTA</i> (0)
<i>GCAT</i> (1)	<i>TGCA</i> (11)
<i>TCGA</i> (32)	<i>ATCG</i> (0)
<i>ATGC</i> (0)	<i>CATG</i> (3)
<i>GATC</i> (45)	<i>CGAT</i> (0)

Interestingly, the above pairs follows the same cyclic permutations, i.e., for example, *A* – *G* – *C* – *T* – *A* – *G* – *C*.

4.1 Distribution of four Letter Combinations

Very less work can be seen in literature having the combinations of *four letters* in *four places*. Most of the work is towards *codon representation* given above. Thus we observe that we $4^4 = 256$ possibilities of writing combinations of four letters in four places. Here below is a configuration 16×16 having all the 256 possibilities.

CCCC	TATA	GTGT	AGAG	CAAT	TCGG	GGTC	ATCA	CTTG	TGCT	GCAA	AAGC	CGGA	TTAC	GACG	ACTT
GTAG	AGGT	CCTA	TACC	GGCA	ATTC	CAGG	TCAT	GCCG	AAAA	CTCT	TGTC	GATT	ACCG	CCAC	TTGA
AGTA	GTCC	TAAG	CCGT	ATGG	GGAT	TCCA	CATC	AACT	GCTG	TGGC	CTAA	ACAC	GAGA	TTTT	CCCG
TAGT	CCAG	AGCC	GTAA	TCTC	CACA	ATAT	GGGG	TGAA	CTGC	AATG	GCCT	TTCC	CGTT	ACGA	GAAC
CTGA	TGAC	GCCG	AATT	CGTG	TTCT	GAAA	ACGC	CCAT	TAGG	GTTC	AGCA	CACC	TCTA	GGGT	ATAG
GCTT	AACG	CTAC	TGGA	GAGC	ACAA	CGCT	TTTG	GTCA	AGTC	CCGG	TAAT	GGAG	ATGT	CATA	TCCC
AAAC	GCGA	TGTT	CTCG	ACCT	GATG	TTGC	CGAA	AGGG	GTAT	TACA	CCTC	ATTA	GGCC	TCAG	CAGT
TGCG	CTTT	AAGA	GCAC	TTAA	CGGC	ACTG	GACT	TATC	CCCA	AGAT	GTGG	TCGT	CAAG	ATCC	GGTA
CGAT	TTGG	GATC	ACCA	CTCC	TGTA	GCGT	AAAG	CAGA	TCAC	GGCG	ATTT	CCTG	TACT	GTAA	AGGC
GACA	ACTC	CGGG	TTAT	GCAG	AAGT	CTTA	TGCC	GGTT	ATCG	CAAC	TCCA	GTGC	AGAA	CCCT	TATG
ACGG	GAAT	TTCA	CGTC	AATA	GCCC	TGAG	CTGT	ATAC	GGGA	TCTT	CACG	AGCT	GTTG	TAGC	CCAA
TTTC	CGCA	ACAT	GAGG	TGGT	CTAG	AACC	GCTA	TCCG	CATT	ATGA	GGAC	TAAA	CCGC	AGTG	GICT
CATG	TCCT	GGAA	ATGC	CCGA	TAAC	GTCG	AGTT	CGCC	TTTA	GAGT	ACAG	CTAT	TGGG	GCTC	AACA
GGGC	ATAA	CACT	TCTG	GTTT	ACCG	CCAC	TAGA	GAAG	ACGT	CGTA	TICC	GCCA	AATC	CTGG	TGAT
ATCT	GGTG	TCGC	CAAA	AGAC	GTGA	TATT	CCCC	ACTA	GACC	TTAG	CCGT	AAGG	GCAT	TGCA	CTTC
TCAA	CAGC	ATTG	GGCT	TACG	CCTT	AGGA	GTAC	TTGT	CGAG	ACCC	GATA	TGTC	CTCA	AAAT	GCCG

The construction of above table is based on the same techniques of the magic of $M_2^{4 \times 4}$. It has the e same properties as of $M_2^{4 \times 4}$. Moreover, the *antiparallel pairs* appearing in the above table are in the same block in each case. They lies in the last eight blocks of order 4×4 .

4.2 Bimagic Squares of Order 16x16

Let us consider now the representations (i) and (ii) of the letters C , A , T and G as given in section 2.1. These representations lead us to following two *bimagic square of order* 16×16 .

4.2.1 First representation

This representation is according (i) given in section 2.1, by choosing $C = 00$, $A = 01$, $T = 10$ and $G = 11$. This we have written in two parts:

Part 1:

00000000	10011001	11101110	01110111	00010110	10001111	11111000	01100001
11100111	01111110	00001001	10010000	11110001	01101000	00011111	10000110
01111001	11100000	10010111	00001110	01101111	11110110	10000001	00011000
10011110	00000111	01110000	11101001	10001000	00010001	01100110	11111111
00101101	10110100	11000011	01011010	00111011	10100010	11010101	01001100
11001010	01010011	00100100	10111101	11011100	01000101	00110010	10101011
01010100	11001101	10111010	00100011	01000010	11011011	10101100	00110101
10110011	00101010	01011101	11000100	10100101	00111100	01001011	11010010
00110110	10101111	11011000	01000001	00100000	10111001	11001110	01010111
11010001	01001000	00111111	10100110	11000111	01011110	00101001	10110000
01001111	11010110	10100001	00111000	01011001	11000000	10110111	00101110
10101000	00110001	01000110	11011111	10111110	00100111	01010000	11001001
00011011	10000010	11110101	01101100	00001101	10010100	11100011	01111010
11111100	01100101	00010010	10001011	11101010	01110011	00000100	10011101
01100010	11111011	10001100	00010101	01110100	11101101	10011010	00000011
10000101	00011100	01101011	11110010	10010011	00001010	01111101	11100100

Part 2:

00101011	10110010	11000101	01011100	00111101	10100100	11010011	01001010
11001100	01010101	00100010	10111011	11011010	01000011	00110100	10101101
01010010	11001011	10111100	00100101	01000100	11011101	10101010	00110011
10110101	00101100	01011011	11000010	10100011	00111010	01001101	11010100
00000110	10011111	11101000	01110001	00010000	10001001	11111110	01100111
11100001	01111000	00001111	10010110	11110111	01101110	00011001	10000000
01111111	11100110	10010001	00001000	01101001	11110000	10000111	00011110
10011000	00000001	01110110	11101111	10001110	00010111	01100000	11111001
00011101	10000100	11110011	01101010	00001011	10010010	11100101	01111100
11111010	01100011	00010100	10001101	11101100	01110101	00000010	10011011
01100100	11111101	10001010	00010011	01110010	11101011	10011100	00000101
10000011	00011010	01101101	11110100	10010101	00001100	01111011	11100010
00110000	10101001	11011110	01000111	00100110	10111111	11001000	01010001
11010111	01001110	00111001	10100000	11000001	01011000	00101111	10110110
01001001	11010000	10100111	00111110	01011111	11000110	10110001	00101000
10101110	00110111	01000000	11011001	10111000	00100001	01010110	11001111

Let us combine the parts 1 and 2 as given below

Part 1	Part 2
--------	--------

This gives us a bimagic square of order 16×16 with $S1^{16 \times 16} := 88888888$ and $S2^{16 \times 16} := 897867554657688$. Each block of order 4×4 is also a magic square with $S1^{4 \times 4} := 22222222$. Square of sum of each term in of each block of order 4×4 is also $S2^{16 \times 16} := 897867554657688$ Here only $S2^{16 \times 16}$ is divisible by 37 i.e, $S2^{16 \times 16} := 897867554657688 = 24266690666424 \times 37$.

4.2.2 Second representation

This representation is according to (ii) given in section 2.1 by choosing $C = 1, A = 1, T = 3$ and $G = 4$.

1111	3232	4343	2424	1223	3144	4431	2312	1334	3413	4122	2241	1442	3321	4214	2133
4324	2443	1132	3211	4412	2331	1244	3123	4141	2222	1313	3434	4233	2114	1421	3342
2432	4311	3224	1143	2344	4423	3112	1231	2213	4134	3441	1322	2121	4242	3333	1414
3243	1124	2411	4332	3131	1212	2323	4444	3422	1341	2234	4113	3314	1433	2142	4221
1342	3421	4114	2233	1434	3313	4222	2141	1123	3244	4331	2412	1211	3132	4443	2324
4133	2214	1321	3442	4241	2122	1413	3334	4312	2431	1144	3223	4424	2343	1232	3111
2221	4142	3433	1314	2113	4234	3341	1422	2444	4323	3212	1131	2332	4411	3124	1243
3414	1333	2242	4121	3322	1441	2134	4213	3231	1112	2423	4344	3143	1224	2311	4432
1423	3344	4231	2112	1311	3432	4143	2224	1242	3121	4414	2333	1134	3213	4322	2441
4212	2131	1444	3323	4124	2243	1332	3411	4433	2314	1221	3142	4341	2422	1113	3234
2144	4223	3312	1431	2232	4111	3424	1343	2321	4442	3133	1214	2413	4334	3241	1122
3331	1412	2123	4244	3443	1324	2211	4132	3114	1233	2342	4421	3222	1141	2434	4313
1234	3113	4422	2341	1142	3221	4314	2433	1411	3332	4243	2124	1323	3444	4131	2212
4441	2322	1213	3134	4333	2414	1121	3242	4224	2143	1432	3311	4112	2231	1344	3423
2313	4434	3141	1222	2421	4342	3233	1114	2132	4211	3324	1443	2244	4123	3412	1331
3122	1241	2334	4413	3214	1133	2442	4321	3343	1424	2111	4232	3431	1312	2223	4144

Here again we have bimagic square of order 16×16 with $S1^{16 \times 16} := 44440$ and $S2^{16 \times 16} := 143634120$. Each block of order 4×4 is also a magic square with $S1^{16 \times 16} := 11110$. Square of sum of each term in of each block of 4×4 is also $S2^{16 \times 16} := 143634120$.

4.2.3 Third representation

Applying the change of base 2 to base 10 (decimal) in first case, i.e.,

$$(abcdefgh)_2 := a \cdot 2^7 + b \cdot 2^6 + c \cdot 2^5 + d \cdot 2^4 + e \cdot 2^3 + f \cdot 2^2 + g \cdot 2^1 + h \cdot 2^0$$

then writing $(abcdefgh)_2 + 1$, we get the *bimagic square of order* 16×16 with sum $S1^{16 \times 16} := 2056$ and $S2^{16 \times 16} := 351576$. Also each block of order 4×4 is a magic square with sum $S1^{4 \times 4} := 514$.

1	154	239	120	23	144	249	98	44	179	198	93	62	165	212	75
232	127	10	145	242	105	32	135	205	86	35	188	219	68	53	174
122	225	152	15	112	247	130	25	83	204	189	38	69	222	171	52
159	8	113	234	137	18	103	256	182	45	92	195	164	59	78	213
46	181	196	91	60	163	214	77	7	160	233	114	17	138	255	104
203	84	37	190	221	70	51	172	226	121	16	151	248	111	26	129
85	206	187	36	67	220	173	54	128	231	146	9	106	241	136	31
180	43	94	197	166	61	76	211	153	2	119	240	143	24	97	250
55	176	217	66	33	186	207	88	30	133	244	107	12	147	230	125
210	73	64	167	200	95	42	177	251	100	21	142	237	118	3	156
80	215	162	57	90	193	184	47	101	254	139	20	115	236	157	6
169	50	71	224	191	40	81	202	132	27	110	245	150	13	124	227
28	131	246	109	14	149	228	123	49	170	223	72	39	192	201	82
253	102	19	140	235	116	5	158	216	79	58	161	194	89	48	183
99	252	141	22	117	238	155	4	74	209	168	63	96	199	178	41
134	29	108	243	148	11	126	229	175	56	65	218	185	34	87	208

According to above three constructions the Reiner [11] table of antiparallelism is given by

Antiparallel A-T, G-C in same orientation	AGCT 01110010 2413 115	CGTA 00111001 1432 58	TACG 10010011 3214 148	CTAG 00100111 1324 40	GCAT 11000110 4123 199	TCGA 10001101 3142 142	ATGC 01101100 2341 109	GATC 11011000 4231 217	SUM 4444444 22220 1028
Antiparallel A-T, G-C in opposite orientation	TAGC 10011100 3241 157	ACGT 01001110 2143 79	GTAC 11100100 4321 229	GCTA 11001001 4132 202	TGCA 10110001 3412 178	ATCG 01100011 2314 100	CATG 00011011 1234 28	CGAT 00110110 1423 55	SUM 4444444 22220 1028

We have the same sum in both the situations, i.e., in the same as well as in the opposite orientation of the genetic letters.

4.3 Hamming Distances and Binomial Coefficients

Here also we shall consider more representations to bring Hamming distances and binomial coefficients. Using the same notations of section 3.4, we have the following table with *Hamming distances* and *binomial coefficients*:

Here we observe the symmetry in elements in each row, each column and each block of order 4×4 . The same symmetry we have in principal diagonals too. This don't happened in case of order 8×8 . Thus there is a straight relationship with the binomial coefficients and Hamming distances, i.e., $0 \rightarrow b^4$, $1 \rightarrow ab^3$, $2 \rightarrow a^2b^2$, $3 \rightarrow a^3b$ and $4 \rightarrow a^4$. Accordingly, we have the following frequency distribution table:

0 b ⁴	4 a ⁴	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b
2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b	0 b ⁴	4 a ⁴	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	1 ab ³	3 a ³ b
3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	4 a ⁴	0 b ⁴
3 a ³ b	1 ab ³	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	4 a ⁴	0 b ⁴	3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²
2 a ² b ²	2 a ² b ²	0 b ⁴	4 a ⁴	1 ab ³	3 a ³ b	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b
2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	0 b ⁴	4 a ⁴	1 ab ³	3 a ³ b	3 a ³ b	1 ab ³
3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	3 a ³ b	1 ab ³	4 a ⁴	0 b ⁴	2 a ² b ²	2 a ² b ²
1 ab ³	3 a ³ b	3 a ³ b	1 ab ³	4 a ⁴	0 b ⁴	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²
2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	0 b ⁴	4 a ⁴	1 ab ³	3 a ³ b	3 a ³ b	1 ab ³
2 a ² b ²	2 a ² b ²	0 b ⁴	4 a ⁴	1 ab ³	3 a ³ b	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b
1 ab ³	3 a ³ b	3 a ³ b	1 ab ³	4 a ⁴	0 b ⁴	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²
3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	3 a ³ b	1 ab ³	4 a ⁴	0 b ⁴	2 a ² b ²	2 a ² b ²
2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b	0 b ⁴	4 a ⁴	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	1 ab ³	3 a ³ b
0 b ⁴	4 a ⁴	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	1 ab ³	3 a ³ b	1 ab ³	3 a ³ b
3 a ³ b	1 ab ³	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	4 a ⁴	0 b ⁴	3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²
3 a ³ b	1 ab ³	3 a ³ b	1 ab ³	2 a ² b ²	2 a ² b ²	2 a ² b ²	2 a ² b ²	3 a ³ b	1 ab ³	1 ab ³	3 a ³ b	2 a ² b ²	2 a ² b ²	4 a ⁴	0 b ⁴

• Frequency distribution

n	Hamming distances	Frequency distributions	Binomial coefficients	Sum
1	0 1	$2^1 = 2$	$a b$	$(a + b)^1$
2	0 1 2	$2^2 = 4$	$a^2 \ 2ab \ b^2$	$(a + b)^2$
3	0 1 2 3	$2^3 = 8$	$a^3 \ 3a^2b \ 3ab^2 \ b^3$	$(a + b)^3$
4	0 1 2 3 4	$2^4 = 16$	$a^4 \ 4a^3b \ 6a^2b^2 \ 4ab^3 \ a^4$	$(a + b)^4$
5	0 1 2 3 4 5	$2^5 = 32$	$a^5 \ 5a^4b \ 10a^3b^2 \ 10a^2b^3 \ 5ab^4 \ b^5$	$(a + b)^5$
6	0 1 2 3 4 5 6	$2^6 = 64$	$a^6 \ 6a^5b \ 15a^4b^2 \ 20a^3b^3 \ 15a^2b^4 \ 6ab^5 \ b^6$	$(a + b)^6$
...

The above table allow us to extend the results for the next values of n. Some studies having combinations of four letters can be seen in [1].

5 Shannon's Entropy and Genetic Tables

The idea of Shannon entropy is well-known in the literature on information theory. It is defined as

$$H(P) = - \sum_{i=1}^n p_i \log p_i$$

where $P = (p_1, p_2, \dots, p_n)$, $p_i > 0$, $\sum_{i=1}^n p_i = 1$ is a set of probability distribution associated with a random variable $X = \{x_1, x_2, \dots, x_n\}$. Applications of Shannon entropy to genetic code can be seen in many works In [13, 14], authors introduce the idea of *genome order index* given by

$$S(P) = \sum_{i=1}^n p_i^2$$

In Information theory the expression $S(P)$ is famous as quadratic entropy. Thus based the magic squares given above we shall calculate *Shannon entropy* and *genome order index* First, we shall transform values in probabilities dividing by sum of each row or column. Here we shall consider only the first case.

5.1 Shannon Entropy of Order 4x4

In section 3.1, we have three different kind of magic squares of order 4×4 . The first one is with binary digits. In this let us divide the each value by the magic sum. This gives us the following probability distributions:

- Probability distributions

0,4500	0,0050	0,0455	0,4995
0,0495	0,4955	0,4550	0,0000
0,5000	0,0450	0,0045	0,4505
0,0005	0,4545	0,4950	0,0500

- Shannon entropy

Based on above probability distributions, let us calculate the values of Shannon entropy. The table below give these values.

				0,3683
0,0400	0,0114	0,0610	0,1506	0,2630
0,0646	0,1511	0,1556	0,0000	0,3713
0,1505	0,0606	0,0106	0,1560	0,3777
0,0015	0,1556	0,1512	0,0650	0,3733
0,2567	0,3788	0,3783	0,3716	0,2667

We observe that the value of Shannon entropy varies from 0,2567 to 0,3777.

5.2 Shannon Entropy of Order 8x8

In sections 3.2.1 and 3.2.2, we have two different kinds of magic squares of order 8×8 . The magic square appearing in section 3.2.2 is bimagic. In both the cases, we considered here below the first one with binary digits. In these cases let us divide the each value by their magic sum. This gives us the following probability distributions:

5.2.1 First case

Here below is a table for probability distributions based on the binary magic square of order 8×8 given in section 3.2.1.

- Probability distribution

0,00000	0,22525	0,24977	0,02498	0,00023	0,22502	0,25000	0,02475
0,24975	0,02500	0,00002	0,22523	0,24998	0,02477	0,00025	0,22500
0,02500	0,24975	0,22523	0,00002	0,02477	0,24998	0,22500	0,00025
0,22525	0,00000	0,02498	0,24977	0,22502	0,00023	0,02475	0,25000
0,00227	0,22748	0,24750	0,02275	0,00250	0,22725	0,24773	0,02252
0,24752	0,02273	0,00225	0,22750	0,24775	0,02250	0,00248	0,22727
0,02273	0,24752	0,22750	0,00225	0,02250	0,24775	0,22727	0,00248
0,22748	0,00227	0,02275	0,24750	0,22725	0,00250	0,02252	0,24773

- **Shannon entropy**

Based on above probability distributions, let us calculate the values of Shannon entropy. The table below give these values.

								0,6766
0,0000	0,1458	0,1505	0,0400	0,0008	0,1458	0,1505	0,0398	0,6732
0,1505	0,0401	0,0001	0,1458	0,1505	0,0398	0,0009	0,1458	0,6734
0,0400	0,1505	0,1458	0,0001	0,0398	0,1505	0,1458	0,0009	0,6734
0,1458	0,0000	0,0400	0,1505	0,1458	0,0008	0,0398	0,1505	0,6732
0,0060	0,1463	0,1501	0,0374	0,0065	0,1462	0,1501	0,0371	0,6797
0,1501	0,0373	0,0060	0,1463	0,1501	0,0371	0,0065	0,1462	0,6796
0,0374	0,1501	0,1463	0,0060	0,0371	0,1501	0,1462	0,0065	0,6796
0,1463	0,0060	0,0374	0,1501	0,1462	0,0065	0,0371	0,1501	0,6797
0,6761	0,6761	0,6761	0,6761	0,6768	0,6768	0,6769	0,6769	0,6763

We observe that the values of Shannon entropy varies from 0,6732 to 0,6797

5.2.2 Second case

In this case we shall deal with bimagic square given in section 3.2.2 with binary digits. The table below is the probability distributions table:

- **Probability distribution**

0,00250	0,22725	0,22502	0,00023	0,02477	0,24998	0,24775	0,02250
0,02475	0,25000	0,24773	0,02252	0,00248	0,22727	0,22500	0,00025
0,00000	0,22525	0,22748	0,00227	0,02273	0,24752	0,24975	0,02500
0,02275	0,24750	0,24977	0,02498	0,00002	0,22523	0,22750	0,00225
0,22523	0,00002	0,00225	0,22750	0,24750	0,02275	0,02498	0,24977
0,24752	0,02273	0,02500	0,24975	0,22525	0,00000	0,00227	0,22748
0,22727	0,00248	0,00025	0,22500	0,25000	0,02475	0,02252	0,24773
0,24998	0,02477	0,02250	0,24775	0,22725	0,00250	0,00023	0,22502

- **Shannon entropy**

								0,6763
0,0065	0,1462	0,1458	0,0008	0,0398	0,1505	0,1501	0,0371	0,6768
0,0398	0,1505	0,1501	0,0371	0,0065	0,1462	0,1458	0,0009	0,6769
0,0000	0,1458	0,1463	0,0060	0,0374	0,1501	0,1505	0,0400	0,6761
0,0374	0,1501	0,1505	0,0400	0,0001	0,1458	0,1463	0,0060	0,6761
0,1458	0,0001	0,0060	0,1463	0,1501	0,0374	0,0400	0,1505	0,6761
0,1501	0,0373	0,0401	0,1505	0,1458	0,0000	0,0060	0,1463	0,6761
0,1462	0,0065	0,0009	0,1458	0,1505	0,0398	0,0371	0,1501	0,6769
0,1505	0,0398	0,0371	0,1501	0,1462	0,0065	0,0008	0,1458	0,6768
0,6763	0,6763	0,6766	0,6766	0,6764	0,6763	0,6766	0,6766	0,6763

- [9] shCherbak, V.I., Arithmetic inside the universal genetic code, *Biosystems*, 70(3)(2003), 187-209.
- [10] Stambuk, N. Universal Metric Properties of Genetic Code, *Croatica Chemica Acta*, **73** (4)(2000), 1123-1139.
- [11] Reiner, B.S., Cyclic Permutations in the Genetic Code, the 4x4 Magic Square and the Antiparallelism of DNA, *J. Theor. Biology*, **110**(1984), 681-690.
- [12] Watson, J. D. and F.H.C. Crick, A structure for deoxyribose nucleic acid. *Nature*, **171**(1953), 737-738
- [13] Zhang, C.T., F. Gao and R. Zhang, Segmentation Algorithm for DNA Sequences, *Physical Review*, **E72**(2005), 041917.
- [14] Zhang, Y., Relations between Shannon Entropy and Genome Order Index in Segmenting DNA Sequences, *Physical Review*, **E79**(2009), 041918.
-